DOCUMENT RESUME

ED 060 863 LI 003 527

AUTHOR	Bhat, U. Narayan; And Others
TITLE	Information Networks: A Probablistic Model for Hierarchical Message Transfer.
INSTITUTION	Southern Methodist Univ., Dallas, Tex. Computer Science/Operations Research Center.
REPORT NO	TR-CP-710023
PUB DATE	Nov 71
NOTE	20p.; (9 References)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Costs; *Electronic Data Processing; *Information Networks; *Models; Theories
IDENTIFIERS	*Information Transfer

ABSTRACT

A strictly hierarchical message transfer scheme requires that a message follow a specified referral path unless finally it is either rejected or filled at any one of the information centers of the network. Thus at each node in the network three decisions can be made: satisfy, reject or refer the message to the succeeding node in the hierarchy. Associating probabilities and costs with each of these decisions, we develop a Markovian model for the total network cost. The mean and variance of total cost are derived. Applicability of the model is discussed by considering the problems related to the estimation of necessary parameters. In particular, a queue theoretic model is developed for estimating response time for a message at an information center. (Author)

INFORMATION NETWORKS: A PROBABILISTIC MODEL
FOR HIERARCHICAL MESSAGE TRANSFER

(Technical Report CP 710023)

U. Narayan Bhat
Richard E. Nance
and
Robert R. Korfhage
Computer Science/Operations Research Center
Institute of Technology
Southern Methodist University
Dallas, Texas 75222

November, 1971

ABSTRACT

A strictly hierarchical message transfer scheme requires that a message follow a specified referral path unless finally it is either rejected or filled at any one of the information centers of the network. Thus at each node in the network three decisions can be made: satisfy, reject or refer the message to the succeeding node in the hierarchy. Associating probabilities and costs with each of these decisions, we devolop a Markovian model for the total network cost. The mean and variance of total cost are derived. Applicability of the model is discussed by considering the problems related to the estimation of necessary parameters. In particular, a queue theoretic model is developed for estimating response time for a message at an information center.

I. Introduction

In Nance, Korfhage and Bhat [1] an information network is defined as a sextuple

$$N = \{U, I, C, \Lambda, f, f'\}$$

where the components of N are defined as below: The entities U, I and C are the set of nodes in the network representing the <u>users</u>, <u>information resources</u> and <u>information centers</u> respectively. We require that with each information center $c \in C$ there be associated a non-empty set $u \in U$ of users, or a non-empty set $i \in I$ of information resources, or both. Λ is the set of directed arcs on $U \cup I \cup C$ where an arc $<v_1,v_2>$ denotes that v_2 is directly accessible from v_1 and where each arc $<v_i,v_j>$ joining nodes of C carries one or both of the labels

m – denoting possible message (request) transfer from v_i to v_j, or

d – denoting possible document (response) transfer from v_i to v_j.

Finally, f and f' are mathematical functions that define the <u>information transfer structure</u> of the network for the message and document transfers respectively.

Using the structural properties of C different types of networks are identified. One of these is the strictly hierarchical network and the purpose of this paper is to study some of its operational characteristics.

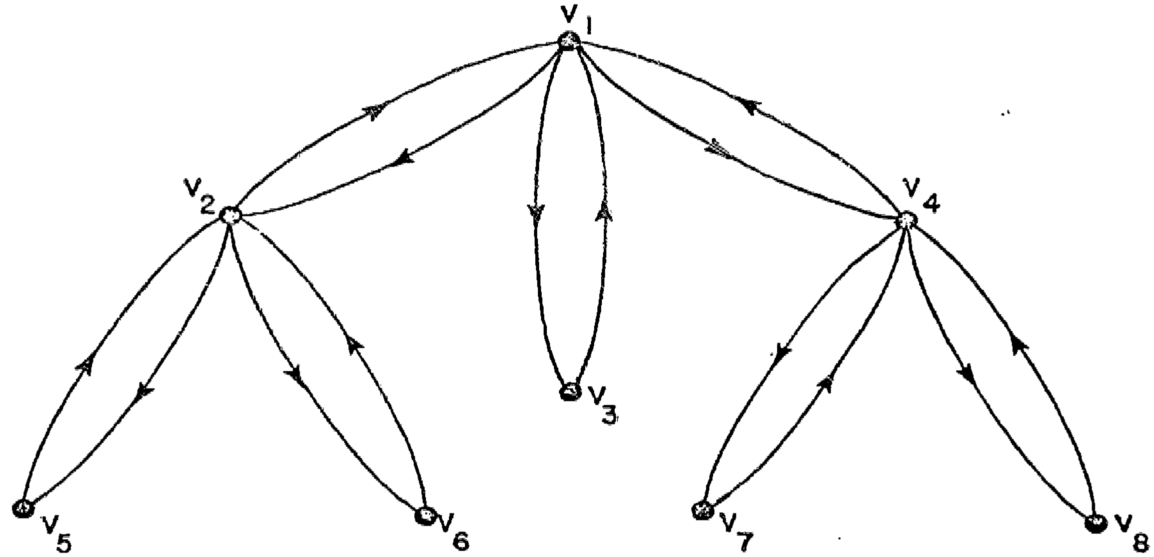
II. Strictly Hierarchical Message Transfer Structure

Let

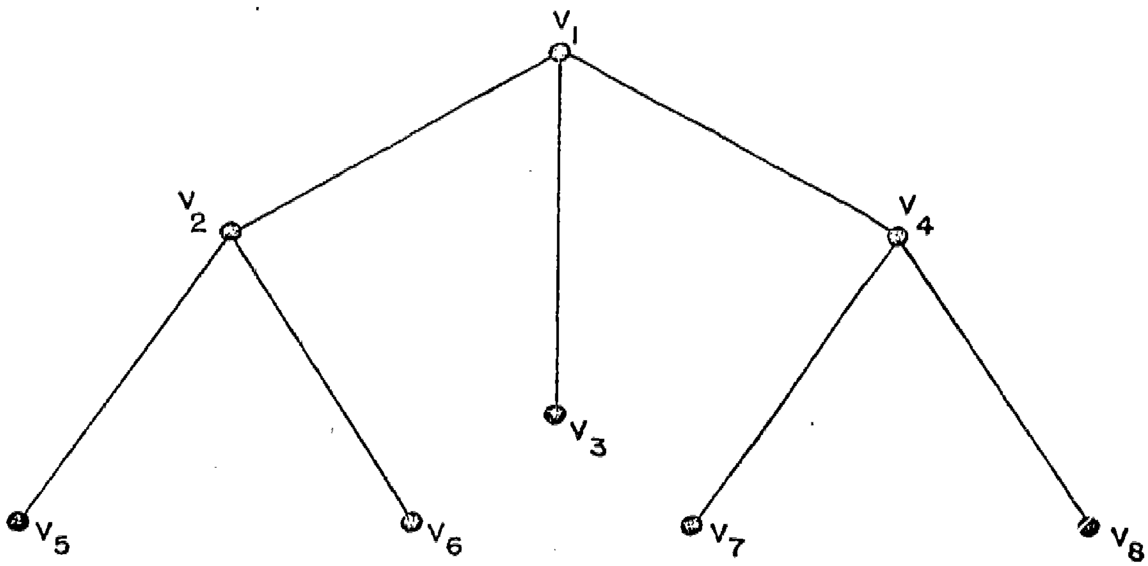
$$G = <C, \text{arcs with label m joining nodes in } C>;$$

then, from [1] we have the following definition.

An information network N is <u>strictly hierarchical</u> if the graph obtained by replacing all 2-cycles in G by an undirected edge is an undirected tree (see [2] for graph theoretic terminology). In such a network any open path joining the root to a leaf is called a <u>limb</u>. The message transfer structure (m.t.s.) in a strictly hierarchical network is illustrated in the following figure.



(a) Stricly Hierarchical Structure.



(b) Replacement of all 2-cycles in G by an undirected edge produces an undirected tree; (V_1,V_2,V_5), (V_1,V_2,V_6), (V_1,V_3), (V_1,V_4,V_7) and (V_1,V_4,V_8) are limbs.

In other words, in a strictly hierarchical network a message follows a unique open path unless finally it is either rejected or filled at any one of the information centers. In view of this, the strictly hiererchical network has one of the most restrictive m.t.s.; and it imposes a unique ordering on the referrals made in the network. Note that other m.t.s. also imply an ordering on the referrals, though not necessarily unique, as demonstrated below.

Consider a network with centers C_1, $C_2,\ldots,C_N$ and a message arriving at an arbitrary center C_{i_o}. If the message cannot be satisfied at C_{i_o}, it can be referred to one or more of the remaining centers for action. For operational effectiveness this referral should take into account the likelihood of satisfying the message at a center and the associated message transfer costs. Let $h_1,h_2,\ldots,h_N$, be the probabilities that a message is satisfied at centers $C_1,C_2,\ldots,C_N$ respectively. Also let c_{ij} be the cost of referring a message from center C_i to center C_j. This cost contains two components: transmission cost and processing cost. Any center involved in the message referral from C_i to C_j will cause a transmission cost. A processing cost for a center is incurred if the decision is to examine the message at that center before referral. For the present discussion we assume only transmission costs for the intermediate centers in the referral paths from C_i to C_j.

Under the above assumptions consider a sequential order of centers for message referrals based on the following scheme.

Let

$$S_{i_o} = \{\mathcal{C}-C_{i_o}\}$$

$$S_{i_{k+1}} = \{\mathcal{C} - \bigcup_{\ell=o}^{k} C_{i_\ell}\}$$

and

$$C_{i_{k+1}} = C_{j_o} \varepsilon S_{i_{k+1}} \;\Big|\; \frac{c_{i_k j_o}}{h_{j_o}} = \min_{j} \frac{c_{i_k j}}{h_j}$$

This implies that the center $C_{i_{k+1}}$ in the referral path of a message originating at C_{i_o} is chosen so as to minimize the cost/probability ratio $c_{i_k j}/h_j$ with regard to the information centers not covered by the message so far. By a slight generalization of [3], this scheme can be shown to result in the least expected cost for the entire message referral process.

This, we feel, is a sufficient justification to consider the strictly hierarchical network for more extensive analysis. In view of the uncertainties involved in the m.t.s. a stochastic model for the referral scheme and the associated costs can be used to seek a better understanding of their economic implications. The effectiveness of the network structure and operation is reflected in the mean and variance of the total cost for the m.t.s. These characteristics also serve as useful criteria for the design of similar networks.

III. A Probabilistic Model for Network Cost

Let $C_1,C_2,\ldots,C_L$ be the L centers constituting a single limb in a strictly hierarchical network such that C_1 is the root and C_L is the leaf. The message referral path is then given by $C_L,C_{L-1},\ldots,C_1$. Associated with each center are three outcomes with respect to each message: rejection, satisfaction or referral to the next center in the limb. For a given message currently at center C_i let $p_{i,-1}$, p_{i0}, and $p_{i,i-1}$ respectively be the probabilities of these outcomes. $(p_{i,-1}+p_{i0}+p_{i,i-1}=1)$.

In constructing a probabilistic model for network cost we shall use the theory of finite Markov chains [4,5]. In order to model the message transfer process in a hierarchical network as a Markov chain, we identify the L centers as L states for the process and add two absorbing states C_{-1} and C_0. These indicate the mode of ultimate disposition of the message. State C_{-1} represents

final rejection of the message and state C_0, a satisfactory disposition. At any time the message transfer process can be considered to be occupying any one of the L+2 states

$$\{C_{-1}, C_0, C_1, C_2, \ldots, C_L\}.$$

It should be noted that once a message reaches either of the states C_{-1} and C_0, it remains there with probability 1. This is represented in the transition probability matrix given below.

$$
P = \begin{array}{rc}
 & \begin{array}{cccccccc} -1 & 0 & 1 & 2 & 3 & \cdots & L-1 & L \end{array} \\
\begin{array}{r} -1 \\ 0 \\ 1 \\ 2 \\ 3 \\ \cdot \\ \cdot \\ L \end{array} &
\left(\begin{array}{cc|cccccc}
1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\hline
p_{1,-1} & p_{10} & 0 & 0 & 0 & \cdots & 0 & 0 \\
p_{2,-1} & p_{20} & p_{21} & 0 & 0 & \cdots & 0 & 0 \\
p_{3,-1} & p_{30} & 0 & p_{32} & 0 & \cdots & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\
p_{L,-1} & p_{L0} & 0 & 0 & 0 & \cdots & p_{L,L-1} & 0
\end{array}\right)
\end{array}
\qquad (3.1)
$$

We represent the matrix P as

$$P = \left(\begin{array}{c|c} I & 0 \\ \hline R & Q \end{array}\right) \qquad (3.2)$$

where I, R and Q correspond to the partitions made in (3.1).

Let c_{ij} ($i=1,2\ldots L$, $j=-1,0,1,2\ldots L$) be the cost associated with a transaction from state C_i to C_j in one step. This may be the actual cost of message transfer or delay encountered in a transaction or a combination of the two. At this stage we shall assume the existence of such costs for the model. Subsequently we explore the possibilities of using response time as part of a cost function.

The cost c_{ij} can be assumed either random or deterministic. When assumed random, we denote its first two moments by

$$E(c_{ij}) = \gamma_{ij} \tag{3.3}$$

and

$$E(c_{ij}^2) = \eta_{ij} \tag{3.4}$$

Let K be the total network cost in a given length of time. This is comprised of costs of messages originating at different centers. Let m_i be the cost associated with a message originating at center C_i, before eventually it is either rejected or filled by one of the centers in the network. Whenever the message is referred to state j, certain costs are incurred based on the possible actions taken. These are

(1) rejection with cost $c_{j,-1}$

(2) satisfaction with cost c_{j0}

(3) referral with cost $c_{j,j-1}$.

Let m_{ij} be the cost associated with such a visit so that

$$m_i = \sum_{j=1}^{i} m_{ij}. \tag{3.5}$$

For m_{ij} we have

$$m_{ij} = 0 \qquad j > i$$

$$m_{ii} = \begin{cases} c_{i,-1} & \text{with probability } p_{i,-1} \\ c_{i0} & \text{with probability } p_{i0} \\ c_{i,i-1} & \text{with probability } p_{i,i-1} \end{cases} \tag{3.6}$$

and

$$m_{ij} = \begin{cases} 0 & \text{with probability } p_{i,-1}+p_{i0} \\ m_{i-1,j} & \text{with probability } p_{i,i-1} \end{cases}$$

$$j < i. \tag{3.7}$$

Let

$$\mu_{ij} = E(m_{ij}) \tag{3.8}$$

$$\sigma_{ij}^{2} = V(m_{ij}) \tag{3.9}$$

for $i,j=1,2,\ldots,L$. The matrices with μ_{ij} and σ_{ij}^{2} as elements are denoted by M and S respectively.

In the following section we use relations (3.6) and (3.7) to derive expressions for μ_{ij} and σ_{ij}^{2} $(i,j=1,2,\ldots,L)$ and the total network cost K.

IV. Mean and Variance of Message Referral Cost

Taking expectations of (3.6) and (3.7) we get

$$\mu_{ii}=p_{i,-1}\gamma_{i,-1}+p_{i0}\gamma_{i0}+p_{i,i-1}\gamma_{i,i-1} \tag{4.1}$$

$$\mu_{ij}=p_{i,i-1}\mu_{i-1,j} \qquad j < i \tag{4.2}$$

Let

$$M_D = \begin{pmatrix} \mu_{11} & & & 0 \\ & \mu_{22} & & \\ & & \ddots & \\ 0 & & & \mu_{LL} \end{pmatrix}. \tag{4.3}$$

Now, expressing (4.1) and (4.2) in matrix notations we get

$$M = M_D + QM \tag{4.4}$$

or

$$M = (I-Q)^{-1}M_D \tag{4.5}$$

provided $(I-Q)^{-1}$ exists.

Clearly

$$(I-Q) = \begin{pmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ -p_{21} & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & -p_{32} & 1 & \cdot & \cdot & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & \cdot & & \cdot & \cdot\, p_{L,L-1} & 1 \end{pmatrix} \qquad (4,6)$$

is non-singular with inverse

$$(I-Q)^{-1} = \begin{pmatrix} 1 & 0 & \cdot & \cdot & 0 \\ p_{21} & 1 & & & \\ p_{32}p_{21} & p_{32} & & & \\ \cdot & & & & \\ \cdot & & & & \\ p_{L,L-1}p_{L-1,L-2}\cdot\cdot p_{21} & \cdot & \cdot & \cdot & 1 \end{pmatrix} \qquad (4.7)$$

thus giving an explicit expression for *M*.

Let n_i be the number of messages originating at center C_i $(i=1,2,\ldots,L)$ during a given period. Then, the expected value of the total network cost K can be expressed as

$$E(K) = \sum_{i=1}^{L} n_i \sum_{j=1}^{L} \mu_{ij}. \qquad (4.8)$$

If n_i is a random variable, we get

$$E(K) = \sum_{i=1}^{L} E(n_i) \sum_{j=1}^{L} \mu_{ij}. \qquad (4.9)$$

To derive the corresponding variances of costs, we square both sides of equations (3.6) and (3.7), to get

$$m_{ii}^{\;2} = \begin{cases} c_{i,-1}^2 & \text{with probability } p_{i,-1} \\ c_{i0}^2 & \text{with probability } p_{i0} \\ c_{i,i-1}^2 & \text{with probability } p_{i,i-1} \end{cases} \tag{4.10}$$

and when $j < i$,

$$m_{ij}^2 = \begin{cases} 0 & \text{with probability } p_{i,-1}+p_{i0} \\ m_{i-1,j}^2 & \text{with probability } p_{i,i-1} \end{cases} \tag{4.11}$$

Taking expectations we can write

$$E(m_{ii}^2) = p_{i,-1}\,\eta_{i,-1} + p_{i0}\eta_{i0} + p_{i,i-1}\eta_{i,i-1} \tag{4.12}$$

and

$$E(m_{ij}^2) = p_{i,i-1}E(m_{i-1,j}^2) \qquad\qquad j < i \tag{4.13}$$

Writing

$$\eta_i = p_{i,-1}\eta_{i,-1} + p_{i0}\eta_{i0} + p_{i,i-1}\eta_{i,i-1} \tag{4.14}$$

and

$$H = \begin{pmatrix} \eta_1 & & & 0 \\ & \eta_2 & & \\ & & \ddots & \\ 0 & & & \eta_L \end{pmatrix} \tag{4.15}$$

We obtain as before

$$||E(m_{ij}^2)|| = (I-Q)^{-1}H \tag{4.16}$$

where we have used $||x_{ij}||$ to denote the matrix with x_{ij} as elements. Let

$$M_2 = \begin{pmatrix} \mu_{11}^{\;2} & \mu_{12}^{\;2} & \cdots & \mu_{1L}^{\;2} \\ \mu_{21}^{\;2} & \mu_{22}^{\;2} & \cdots & \mu_{2L}^{\;2} \\ \vdots & \vdots & & \vdots \\ \mu_{L1}^{\;2} & \mu_{L2}^{\;2} & \cdots & \mu_{LL}^{\;2} \end{pmatrix} \tag{4.17}$$

Then we can write

$$S = (I-Q)^{-1}H - M_2. \tag{4.18}$$

To derive the variance of the total cost K we also need

$$E[(\sum_{j=1}^{i} m_{ij})^2].$$

We have

$$m_i = \sum_{j=1}^{i} m_{ij}. \tag{4.19}$$

From (3.6) and (3.7) we get

$$m_i = \begin{cases} c_{i,-1} & \text{with probability } p_{i,-1} \\ c_{i0} & \text{with probability } p_{i0} \\ c_{i,i-1}+m_{i-1} & \text{with probability } p_{i,i-1} \end{cases} \tag{4.20}$$

Squaring both sides of (4.20) and taking expectations we get

$$\begin{aligned} E(m_i^2) &= p_{i,-1}\eta_{i,-1} + p_{i0}\eta_{i0} + p_{i,i-1}\eta_{i,i-1} \\ &\quad + p_{i,i-1}E(m_{i-1}^2) + 2p_{i,i-1}E(c_{i,i-1}m_{i-1}). \end{aligned} \tag{4.21}$$

In the hierarchical network structure $c_{i,i-1}$ and m_{i-1} can be assumed to be independent random variables; hence

$$E(c_{i,i-1}m_{i-1}) = \gamma_{i,i-1}\mu_{i-1} \tag{4.22}$$

where we have written $\sum_{j=1}^{L} \mu_{ij} = \mu_i = E(m_i)$.

Let

$$\Gamma = \begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & 0 & 0 \\ p_{21}\gamma_{21} & 0 & 0 & \cdot & \cdot & 0 & 0 \\ 0 & p_{32}\gamma_{32} & 0 & \cdot & \cdot & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & & & p_{L,L-1}\gamma_{L,L-1} & 0 \end{pmatrix} \tag{4.23}$$

In matrix notation (4.24) can then be written as

$$\begin{pmatrix} E(m_1^2) \\ E(m_2^2) \\ \vdots \\ E(m_L^2) \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_L \end{pmatrix} + Q \begin{pmatrix} E(m_1^2) \\ E(m_2^2) \\ \vdots \\ E(m_L^2) \end{pmatrix} + 2\Gamma \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_L \end{pmatrix} \tag{4.24}$$

which on re-arrangement gives

$$\begin{pmatrix} E(m_1^2) \\ E(m_2^2) \\ \vdots \\ E(m_L^2) \end{pmatrix} = (I-Q)^{-1} \left[\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_L \end{pmatrix} + 2\Gamma \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_L \end{pmatrix} \right] \tag{4.25}$$

Let n_i be the number of messages originating at center C_i, as assumed before. We also assume that $E(n_i)$ and $E(n_i^2)$ exist and are known. Let $m_i^{(1)}, m_i^{(2)}, \ldots, m_i^{(n_i)}$ be the costs associated with these n_i messages. Total cost of all the messages originating at center C_i is given by

$$K_i = m_i^{(1)} + m_i^{(2)} \ldots + m_i^{(n_i)} .$$

We assume that costs $m_i^{(r)}$ $(r=1,2,\ldots,n_i)$ are independent of the number of messages n_i. The variance of this sum is given by

$$V\left(\sum_{r=1}^{n_i} m_i^{(r)}\right) = V(n_i)\,[E(m_i)]^2 + E(n_i)V(m_i). \tag{4.26}$$

When n_i's are deterministic, we have

$$V\left(\sum_{r=1}^{n_i} m_i^{(r)}\right) = n_i V(m_i). \tag{4.27}$$

For the total cost K, we have

$$V(K) = V\left(\sum_{r=1}^{n_1} m_1^{(r)} + \sum_{r=1}^{n_2} m_2^{(r)} + \ldots + \sum_{r=1}^{n_L} m_L^{(r)}\right)$$

$$= \sum_{i=1}^{L} V\left(\sum_{r=1}^{n_i} m_i^{(r)}\right) + 2\sum_{i<j}\sum \operatorname{cov}\left(\sum_{r=1}^{n_i} m_i^{(r)}, \sum_{k=1}^{n_j} m_j^{(k)}\right). \tag{4.28}$$

When n_i, n_j are random the covariance term on (4.28) is very involved. When they are constant, however,

$$\operatorname{Cov}\left(\sum_{r=1}^{n_i} m_i^{(r)}, \sum_{k=1}^{n_j} m_j^{(k)}\right) = n_i n_j \sum_{\ell=1}^{L} q_{i\ell} q_{j\ell} (\eta_\ell - \mu_{\ell\ell}^2) \tag{4.29}$$

where

$$q_{i\ell} = \begin{cases} p_{i,i-1} p_{i-1,i-2} \cdots p_{\ell+1,\ell} , & \ell < i \\ 0 \text{ otherwise.} & \end{cases}$$

Thus we get

$$V(K) = \sum_{i=1}^{L} n_i V(m_i) + 2\sum_{i<j}\sum n_i n_j \sum_{\ell=i}^{L} q_{i\ell} q_{j\ell} (\eta_\ell - \mu_{\ell\ell}^2). \tag{4.30}$$

V. Use of Model for Evaluation and Design

Introduction of a mathematical model is only of academic interest unless its usefulness in solving real world problems is explained. In this section we discuss the requirements for applying the model developed in the preceding two sections.

The Markovian model can be used in two different situations: (1) evaluation of existing networks and (2) development of design criteria for new networks.

The decision to apply the Markovian model to an existing network introduces the problem of parameter estimation. The basic parameters in our model are

the probability elements of the matrix P in (3.1). Maximum likelihood estimates of these probabilities are given by the fraction of messages referred from one center to another [6]. That is, if n_i is the number of messages received by center C_i during an observation period, and n_{ij} of these are referred to center C_j, then the estimate $\hat{p}_{ij}$ of the probability p_{ij} is given by

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i} . \tag{5.1}$$

For an operating network, one can derive means and variances of costs incurred by the message referral process. No theoretical forms for distributions need be assumed; empirical distributions can be used. Also different forms of cost functions can be used to describe the situation. Once these parameters have been estimated, the expected message cost can be obtained using formulas developed in section IV and comparisons can be made between two or more existing networks regarding their cost and effectiveness.

The extensive data collection necessary for model construction is provided by the statistical reporting systems currently being implemented on digital computers. An example is the TALON Medical Library Network's TRIPS System [7].

Applying the Markovian model to network design poses several problems related to parameter estimation, some of which can be handled through standard techniques. Sufficient data may exist for the estimation of parameters:

(1) relating to message arrivals,
(2) describing the information resources,
(3) providing the message referral probabilities,
(4) indicating alternative message transfer modes, and
(5) furnishing costs involved throughout the network.

One estimate that is not easily obtained from data is the response time for network requests. We must use a mathematical model to estimate response time as a function of message arrival and service rates at individual centers of the

network.

The mathematical model proposed for this purpose is a simple model from queueing theory [8, 9]. At each center the arriving messages are analogous to customers in a queueing system and the processing of messages to the service function. With this similarity in mind, we make the following assumptions:

(1) At center C_i, users initiate request for information at the rate of λ_{1i} per unit time.

(2) Center C_i also receives requests from center C_{i+1} at the rate of $\lambda_{i+1}p_{i+1,i}$ per unit time. Let

$$\lambda_i = \lambda_{1i} + \lambda_{i+1}p_{i+1,i} \tag{5.2}$$

Thus λ_i is the combined arrival rate at center C_i.

(3) The combined arrival process has the characteristics of a Poisson process with parameter λ_i; i.e., if $A(t)$ is the number of messages arriving during an interval of length t, then

$$\Pr\{A(t)=n\} = e^{-\lambda_i t}\frac{(\lambda_i t)^n}{n!} \quad (n=0,1,2,\ldots). \tag{5.3}$$

(4) The message processing times at center C_i are independent and identically distributed random variables with mean b_i^1 and second moment b_i^2 $(i=1,2,\ldots,L)$.

(5) Processing times are independent of the message arrival process and the number of messages waiting to be processed.

Let ρ_i be the utilization factor for center C_i defined by

$$\rho_i = \frac{\text{Rate of arrival of messages}}{\text{Rate of processing}} = \lambda_i b_i^1 . \tag{5.4}$$

Let R_i be the expected response time at center C_i which is defined as the time interval from the arrival of a message until its disposition at that center.

Based on the assumptions above, we can identify the message processing at a center with the operation of a single server queueing system with similar characteristics. Then the response time is given by [equation (1.196) of [9]]

$$R_i = b_i^1 + \frac{\lambda_i b_i^2}{2(1-\rho_i)}. \tag{5.5}$$

In (5.5) estimates of parameters can be obtained through standard techniques; therefore, estimates of R_i $(i=1,2,\ldots,L)$ can be derived. These estimates can be used in the determination of the expected costs γ_{ij} of (3.3) either directly or in conjunction with other factors such as message transfer costs.

VI. Summary and Discussion

Taking into account the uncertainties associated with the decisions made at each center of a strictly hierarchical network, a probabilistic model is developed for the network cost. Expressions are given for the mean and variance of the cost, and methods are suggested for the estimation of model parameters required for application.

The strictly hierarchical network is very restrictive as evidenced by the flexibility measure developed in [1] as well as the transition probability matrix presented in (3.1). Even though the discussion in section II justifies the use of the strictly hierarchical structure in many practical situations, many types of network operations exist which do not belong to this class [1, section III]. In such cases the transition probabilities of the message referral schemes are usually non-stationary; thus further research is needed for their analysis.

In section II, a sequential order of centers for message referrals is presented assuming that the referral cost c_{ij} between centers C_i and C_j are independent of the centers through which the message is referred. A more realistic approach is to consider the probabilities of satisfying the request at intermediate centers.

The response time estimate is derived in (5.5) under very restrictive assumptions. In assumption (3) of section V, the combined arrival process is assumed to have the characteristics of a Poisson process. If the user requests at each center follow a Poisson process and the processing times at each center can be represented by independent and identically distributed random variables with the negative exponential density function

$$f(x) = \begin{cases} \mu e^{-\mu x} & x>0 \\ 0 & x\leq 0 \end{cases}, \tag{6.1}$$

then, in the long run, the Poisson assumption is justified. Note that we have not assumed any specific form for the processing time distribution in assumption (4). This reflects our belief that the mixture of different types of arrivals at a center may justify the Poisson assumption even when the processing time is not exponentially distributed. This contention needs to be tested through data collected from network operation.

In assumption (4) of section V a single set of first and second moments has been assumed for processing times at a center irrespective of the nature of message disposition. Use of different sets of parameters for different types of messages should present no problems in the extension of this model.

In the above discussion we emphasize that the present investigation solves only some aspects of the general problem. We believe that further research on the remaining aspects should provide a strong theoretical base in the analysis and design of information networks.

References

1. Richard E. Nance, Robert R. Korfhage and U. Narayan Bhat, "Information Networks: Definitions and Message Transfer Models", Technical Report CP 710011, Computer Science/Operations Research Center, Institute of Technology, Southern Methodist University, Dallas, Texas, July, 1971.

2. Harary, F., Graph Theory, Addison-Wesley, 1969.

3. Baker, Norman R., "Optimal User Search Sequences and Implications for Information Systems Operation," American Documentation, 20 (3), 1969, 203-212.

4. Kemeny, J. G. and Snell, J. L., Finite Markov Chains, D. Van Nostrond Co., Princeton, New Jersey, 1960.

5. Bhat, U. Narayan, Elements of Applied Stochastic Processes, John Wiley & Sons, New York, 1972.

6. Anderson, T. W. and Goodman, L. A., "Statistical Inference about Markov Chains," Annals of Math. Stat. 28 (1957), 89-110.

7. Nance, R. E., Wickham, W. K. and Duggan, M., "A Computer System for Effective Management of a Medical Library Network: An Overview," J. of Library Automation, December, 1971.

8. Saaty, T. L., Elements of Queueing Theory, McGraw-Hill, New York, 1961

9. Prabhu, N. U., Queues and Inventories, John Wiley & Sons, New York, 1965.

COMPUTER SCIENCE /
OPERATIONS RESEARCH CENTER
SMU INSTITUTE OF TECHNOLOGY
DALLAS, TEXAS 75222